

DOCUMENT RESUME

ED 401 282

TM 025 258

AUTHOR McGaghie, William C.
TITLE Comparison of Knowledge Structures with the Pathfinder Scaling Algorithm.
PUB DATE Apr 96
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Algorithms; *Cognitive Structures; Comparative Analysis; Higher Education; *Instructional Effectiveness; Medical Education; Medical School Faculty; *Medical Students; Multidimensional Scaling; Physicians; Physiology; *Scaling
IDENTIFIERS *Concept Networks; Experts; Similarity (Concept)

ABSTRACT

The cognitive structure of 13 concepts in pulmonary physiology was explored among 112 first-year medical students and among 32 faculty members in three different expertise groups in a knowledge representation study. Purposes were to assess the degree of agreement among faculty members, map students' concept structures, and compare the similarity of student concepts with those of instructors. It was hypothesized that, as a consequence of instruction, students' concept networks would approximate the structure of concept networks produced by faculty experts. The Pathfinder scaling algorithm was used to map and compare student and faculty concept structures. Pathfinder uses pairwise judgments of similarity between a set of concepts to produce a network or graph and to focus on the local relationships among concepts, in contrast to multidimensional scaling, which seems to capture more global information about the concept space. Data show that the ways in which medical experts conceptually organized the concepts were not necessarily consistent. Similarity ratings for the three expert groups (internists, anesthesiologists, and physiologists) were modest, and comparisons between groups were weaker than comparisons within groups. Data do support the hypothesis that students' Pathfinder networks would begin to approximate the structure of the concept networks produced by faculty as a consequence of instruction, although the choice of which faculty network is best remains uncertain. (Contains 1 table, 1 figure, and 19 references.) (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

Comparison of Knowledge Structures with the Pathfinder Scaling Algorithm

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to
improve reproduction quality.

• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

WILLIAM MCGAGHIE

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

William C. McGaghie
Office of Medical Education
Northwestern University Medical School

April, 1996

Presented at the 1996 annual meeting of the American Educational Research Association, New
York, N.Y. Session No. 12.12, Alternative Approaches to Research on the Cognitive
Structure of Concepts

Comparison of Knowledge Structures with the Pathfinder Scaling Algorithm

This report has three sections. The first section provides a very short distillation of a framework for organizing techniques to represent expert knowledge. The framework was created and published by J. R. Olson and K. J. Biolsi (1991) and is used here for convenience and clarity. Section two describes and interprets the results of a study on the cognitive structure of 13 concepts in pulmonary physiology. This research involved freshmen medical students and faculty members at the Northwestern University Medical School. The study used the Pathfinder scaling algorithm to create conceptual maps for students and faculty, graphs which were used to compare student and faculty knowledge structures. The third section sets forth several "next steps" for this research and development program including conceptual and methodological refinements formed by experience and applications to instructional problems.

Techniques for Representing Expert Knowledge

Olson and Biolsi (1991) provide a framework to classify research studies which address the representation of expert knowledge. In brief, these scholars first distinguished direct vs. indirect techniques of knowledge representation. Next, they describe progressive steps in knowledge representation research. The steps are (a) eliciting behavior from respondents, (b) summarizing behavioral responses as data, (c) performing data analyses, and (d) displaying regularities in the output from the data analyses.

Direct vs. indirect techniques. Olson and Biolsi (1991) describe direct knowledge representation methods as those which rely chiefly, but not exclusively, on verbal reports which are transcribed and studied using propositional or discourse analyses. Such methods include interviews, thinking aloud protocols, and behavioral observations. Olson and Biolsi also allow for various forms of drawing and card sorting as direct methods of knowledge representation. Research by scholars on this symposium panel is prominent for its use of direct methods. Examples include studies done by Georges Bordage and his research group (Bordage, 1994; Bordage and Lemieux, (1991); Katherine Edmonson's (1994) work involving concept maps; and research by Patel and colleagues (1993) which has frequently employed thinking aloud protocols.

Indirect methods contrast with direct techniques because respondents usually perform a rating or judgment task which yields raw data. These data are subsequently analyzed using one or more methods such as multidimensional scaling (MDS), various forms of clustering or tree fitting, or a general weighted network to yield a representational structure of concepts under study. Work my colleagues and I have recently reported (McGaghie, Boerger, McCrimmon, and Ravitch, 1994, 1996) has employed an indirect method, Pathfinder, to graph the structure of one set of concepts embedded in a basic medical science.

Steps in knowledge representation research. The Olson and Biolsi framework identifies four steps involved in any knowledge representation study, independent of the specific research method employed. The steps are:

1. Elicit behavior from respondents. Examples of direct methods to elicit behavior include interviews and thinking aloud protocols. An indirect method to elicit behavior for knowledge representation research is to have respondents judge the pairwise similarities of a set of concepts. These pairwise judgments are typically cast on a seven or nine-point scale, i.e., 1 = completely unrelated, 7 = highly related.
2. Summarize behavioral responses in a manner that permits qualitative or quantitative data analysis. In my own research this has involved summarizing judgment data as lower left proximity half-matrices (deBliek, McGaghie, and Donohue, 1984; McGaghie, Boerger, McCrimmon, and Ravitch, 1994, 1996).
3. Data analysis using one or more of a variety of quantitative or qualitative methods. My own work has employed several approaches to data analysis including multidimensional scaling (deBliek, McGaghie, and Donohue, 1984) and Pathfinder scaling (McGaghie, Boerger, McCrimmon, and Ravitch, 1994, 1996).
4. Display regularities in the output from the data analyses. Illustrations include qualitative concepts maps reported by Katherine Edmonson (1994), search path mapping as performed by Ronald Stevens and colleagues (Stevens, 1991; Stevens, Lopo, and Wang, 1996), descriptions of elaborated thinking presented by Georges Bordage (1994), and Patel's (1993) descriptions of the results of propositional analyses of thinking aloud protocols.

The next section of this paper gives a description of a knowledge representation study conducted at the Northwestern University Medical School in the context of a first year physiology course.

Structure of Pulmonary Physiology Concepts

A recent study at Northwestern University Medical School explored the structure of 13 concepts in pulmonary physiology among 112 first-year medical students and among faculty members, individually and in groups (McGaghie, Boerger, McCrimmon, and Ravitch, 1994, 1996). The goals of this research were to (a) assess the degree of agreement among faculty members about the cognitive structure of the concepts, (b) map students' concept structures before and after a period of focused instruction, and (c) compare the similarity of students' concept structures (especially postinstruction) with those of faculty experts. We hypothesized that as a consequence of instruction, students' concept networks would approximate the structure of concept networks produced by faculty experts, a finding that has been demonstrated elsewhere (Goldsmith, Johnson, and Acton, 1991).

Pathfinder description. The approach we used to map and compare student and faculty concept structures is the Pathfinder scaling algorithm (Johnson, Goldsmith, and Teague, 1995; Schvaneveldt, Durso, and Dearholt, 1989). Pathfinder uses pairwise judgments of similarity between a set of concepts to produce a network or graph that shows the latent structure of a conceptual domain for individuals or groups. Pathfinder networks represent concepts as nodes and distances between concepts as network links. The algorithm takes proximity values as input and yields a network with shortest-path link distances as output. Pathfinder networks of concept structures differ from those derived by MDS because Pathfinder focuses on the local relationships among concepts, whereas MDS seems to capture more global information about the concept space (Gonzalvo, Cañas, and Bajo, 1994).

The configural similarity of two Pathfinder networks having a common set of nodes can be measured by the index PFC', which indicates how many links between concept pairs they have in common (Goldsmith and Davenport, 1990). For example, consider two networks, each with four nodes (A, B, C, and D). Network 1 has three links (A-B, A-C, and A-D), and Network 2 has three links (A-B, A-D, and B-C). the PFC' for these two networks would be .50, the ratio of the number of links the networks have in common (two: A-B and A-D) to the four links in the network. Another index, coherence, indicates the "internal consistency" of a network by comparing direct and indirect measures of the similarity of each concept pair. For each pair of items, an indirect index of relatedness is computed by

correlating the proximities between the items and all other items. Coherence is then computed by correlating the original proximity data with the indirect measures. The same value can be computed for a network derived from aggregate data, indicating the internal consistency of these networks. Finally, the raw proximity data from a number of subjects can be averaged to yield an aggregate data set and network representation.

Data collection. Faculty members of a first-year medical physiology course identified 13 pulmonary concepts medical students are expected to master. The concepts are alveolar ventilation, compliance, dead space, diffusion, gas exchange, hydrostatic pressure, intrapleural pressure, lung perfusion, partial pressure gradient, resistance to flow, shunt flow, solubility, and surface area. A questionnaire was constructed that presented all possible $n(n-1)/2 = 78$ pairs of the concepts randomized for presentation order (left to right) and sequence (1-78). Respondents were instructed to judge the similarity or relatedness of the two concepts in each pair on a scale of 1 (completely unrelated) to 7 (highly related). Concept pairs judged to be highly related received a 6 or 7. Concepts judged unrelated received a 1 or 2. Data collection for the student and faculty respondents took approximately 30 minutes.

Data were obtained from the 112 medical students immediately before and after an educational unit on pulmonary physiology that lasted three weeks. The students also took a 54 item multiple-choice examination that was administered in class at the end of the unit. Data were also collected from three groups of medical faculty experts: (a) 11 internists, (b) 17 anesthesiologists, and (c) 3 physiologists. the physiology course professor also completed the questionnaire.

Data analysis. Data analysis was done in five steps. First, separate Pathfinder analyses were performed to create individual concept networks for each respondent. This yielded 31 faculty networks and 224 student networks (112 pre and post instruction). Second, proximity data for each group of experts were averaged and a network structure was derived for each group using the aggregate data. An index of coherence was computed for each aggregate network. Within and between group similarities were also assessed. Third, similarity measures were computed comparing each student network with the aggregate expert networks and the course professor's network. Fourth, similarity data were averaged to yield similarity scores which indicate how similar student networks were to experts' networks, pre

and post instruction. Fifth, student-expert similarity scores were correlated with multiple-choice examination scores as an index of learning accomplished in the educational unit.

Results. Graphs produced by the individual Pathfinder analyses were striking for their idiosyncratic qualities. Inspection of the 31 faculty networks revealed marked discrepancies in the experts' concept organizations within and between the three groups. For the three groups of experts the average configural similarity of their Pathfinder networks was modest, yet positive: internists = .29 (s.d. = .07), physiologists = .34 (s.d. = .07), and anesthesiologists = .33 (s.d. = .12). Much less similarity was observed across the expert groups: internists - physiologists = .28 (s.d. = .08), physiologists - anesthesiologists = .25 (s.d. = .08), and internists - anesthesiologists = .22 (s.d. = .07). A comparison of within and between-group similarity results showed this difference to be significant ($t = 8.74$, $p < .001$).

Results of Pathfinder analyses performed on two of the three sets of group data yielded interpretable (coherent) results (McGaghie, Boerger, McCrimmon, and Ravitch, 1994). For the internists, the graph indicated the centrality of alveolar ventilation and lung perfusion among the concepts and to a slightly lesser extent to diffusion and gas exchange. There is strong coherence for the internists' aggregate data (.61). The Pathfinder conceptual network for the three physiologists gives a different picture. Here, gas exchange was perceived as the central concept with lung perfusion and several others providing lower degrees of conceptual focus. The coherence index for the physiologists' aggregate data is acceptable (.48), yet lower than that obtained for the general internists. However, an aggregate Pathfinder network could not be presented for the anesthesiologists' judgments because the index of coherence for their data was near zero (.07). Such low coherence rules out a meaningful interpretation of this network structure.

Results from the Pathfinder analysis of the concept judgment made by the individual course professor are shown in Figure 1. The network shows that in this person's judgment the central concept is gas exchange (seven node links), and to a lesser extent, diffusion (four node links), and lung perfusion (three node links). For this professor, the concepts shunt flow and intrapleural pressure are distal items within the context of the 13 concept set.

Figure 1 here

As shown in Table 1, evaluation of the similarity of student and expert networks reveals an increase in similarity between students' networks and all experts' networks after completion of the three-week unit, with the smallest increase for the students' similarity with internists and the greatest change for the students' similarity with the group of three physiologists. These similarity scores were then used to predict multiple-choice examination performance. These results are also shown in Table 1. The reliability (KR-20) of the multiple-choice examination was .54, setting an upper limit for these correlations. The student-expert similarity scores using postinstruction student networks showed statistically significant (**but very modest**) correlations with final examination scores for the groups of internists, physiologists, and the individual course professor.

Table 1 here

These results demonstrate that as a consequence of instruction, the conceptual networks of students become more similar to those of experts in the field as well as to those of their instructor. The similarity index can also be used to predict performance on a traditional, multiple-choice assessment. Thus, they appear to reflect changes in the way students come to think about and organize a body of knowledge.

Discussion. The reported data show that the ways in which medical experts conceptually organize concepts in pulmonary physiology are not necessarily consistent. Similarity ratings for all three expert groups were modest, and comparisons between groups were significantly weaker than those within groups. In addition, of the three expert groups studied, only the general internists produced a concept network having strong coherence (internal consistency). The network structure derived from the physiologists was interpretable, with modest yet acceptable coherence. By contrast, the network produced by the anesthesiologists' aggregate data lacked enough coherence to be interpretable.

The use of experts' network structures of concepts as a "gold standard" for comparison of student networks may pose a problem. The problem exists both for using an individual expert's concept network as a "gold standard" for student performance (Goldsmith, Johnson, and Acton, 1991) and for aggregate network structures (Acton, Johnson, and Goldsmith, 1994). Judgments about the validity of a concept network, either for descriptive research or for student evaluation, depend on the individual or group whose perceptions underlie the network. While the Northwestern data and data reported by other investigators (Goldsmith, Johnson, and Acton, 1991; Gonzalvo, Cañas, and Bajo, 1994) support the hypothesis that students' Pathfinder networks will begin to approximate the structure of concept networks produced by faculty as a consequence of instruction, the choice of which faculty network is best remains uncertain (Johnson, Goldsmith, and Teague, 1995).

These and other data also support the idea that knowledge is constructed by individuals and does not exist as an independent, objective entity (Schwandt, 1994). This view accounts for the absence of perfect coherence indexes for the aggregate faculty networks used as "gold standards" of comparison for student performance. The Northwestern data also support observations about professional workers made by Schön (1993) concerning "knowledge in action" and by Wagner (1991) concerning "tacit knowledge." These scholars argue that experienced professionals recognize, comprehend, frame, and address professional problems in unique yet efficient ways, especially when problems are complex and uncertain. Thus it is not surprising that even presumably homogeneous groups of faculty experts derived concept structures having much variation.

Finally, another key validity issue concerns the use of students' multiple-choice examination scores as an external criterion against which to correlate student-expert similarity

indexes. This can result in circular reasoning, as described by Johnson, Goldsmith, and Teague (1995):

"The external criteria used in this work, either classroom exam performance or performance on ...ACT subtest, poses a potential problem. On the one hand, we claim to be designing an instrument that is a better assessment of structural knowledge than conventional classroom exams. However, the classroom exam, the very measure we are criticizing, is used as the external criterion to argue for the validity of our methods" (pp. 243-244).

Clearly, there is no shortage of intellectual and technical issues to address as research involving Pathfinder scaling and other approaches to knowledge representation moves ahead.

Next Steps

At Northwestern, our research group plans to follow the suggestions of Johnson, Goldsmith, and Teague (1995) regarding "Structurally Driven Training." Such work will progress in several stages.

- Identify a set of educational concepts students are expected to acquire or master.
- Acquire pairwise similarity judgments of the concepts from faculty experts, individually and in groups.
- Acquire pairwise similarity judgments of the concepts from students before instruction.
- Perform Pathfinder scaling of the faculty and student data.
- Publicly present and use faculty concept networks as "advance organizers" for students early in the instructional period. In open discussion different (medical) faculty (e.g., physiologists, general internists, subspecialty pulmonologists, anesthesiologist) would attempt to explain why they intellectually organize concepts differently. They would also attempt to account for the origins of the expected differences.
- Students would actively compare their individual networks with those of faculty experts.

- Student concept networks would also be divided after the instructional period to assess the extent to which they begin to approximate the thinking of faculty. Of course, **which** faculty members the students tend to approximate will also be of great interest.

This and other evolving projects will occupy the thought and energy of Northwestern investigators interested in using the Pathfinder method of knowledge representation to study individual differences in cognition. We recognize, of course, the variety and richness of alternative approaches to research on the cognitive structure of concepts.

REFERENCES

- Acton, W. H., Johnson, P. J., and Goldsmith, T. E. (1994). Structural Knowledge Assessment: Comparison of Referent Structures. Journal of Educational Psychology, 86, 303-311.
- Arocha, J.F., Patel, V.L., and Patel, Y.C. (1993). Hypothesis Generation and the Coordination of Theory and Evidence in Novice Diagnostic Reasoning. Medical Decision Making, 13, 198-211.
- Bordage, G. (1994). Elaborated Knowledge: A Key to Successful Diagnostic Thinking. Academic Medicine, 69, 883-885.
- Bordage, G., and Lemieux, M. (1991). Semantic Structures and Diagnostic Thinking of Experts and Novices. In Proceedings of the Thirtieth Annual Conference on Research in Medical Education. Academic Medicine, 66, S70-S72 (September Supplement).
- de Blik, R., McGaghie, W. C., and Donohue, J.F. (1984). Representation of Clinical Case Cues: A Multidimensional Scaling Demonstration. In Proceedings of the Twenty-Third Annual Conference on Research in Medical Education. S. Kerby, Compiler. Washington, D.C.: Association of American Medical Colleges.
- Edmondson, K. M. (1994). Concept Maps and the Development of Cases for Problem-based Learning. Academic Medicine, 69, 108-110.
- Goldsmith, T. E. and Davenport, D. M. (1990). Assessing Structural Similarity of Graphs. In R. W. Schvaneveldt (Ed.), Pathfinder Associative Networks: Studies in Knowledge Organization. Norwood, N.J.: Ablex Publishing Co., 1990, Pp. 75-87.
- Goldsmith, T. E., Johnson, P. J., and Acton, W. H. (1991). Assessing Structural Knowledge. Journal of Educational Psychology, 83, 88-96.

- Gonzalvo, P., Cañas, J. J., and Bajo, M-T. (1994). Structural Representations in Knowledge Acquisition. Journal of Educational Psychology, 86, 601-616.
- Johnson, P. J., Goldsmith, T. E., and Teague, K. W. (1995). Similarity, Structure, and Knowledge: A Representational Approach to Assessment. In P. D. Nichols, S. F. Chipman, and R. L. Brennan (Eds.), Cognitively Diagnostic Assessment. Hillsdale, N.J.: Lawrence Erlbaum Associates, Pp. 221-249.
- McGaghie, W. C., Boerger, R. L., McCrimmon, D. R., and Ravitch, M. M. (1994). Agreement Among Medical Experts About the Structure of Concepts in Pulmonary Physiology. Academic Medicine, 69(10), S78-S80, (October Supplement).
- McGaghie, W. C., Boerger, R. L., McCrimmon, D. R., and Ravitch, M. M. (1996). Learning Pulmonary Physiology: Comparison of Student and Faculty Knowledge Structures. Academic Medicine, 71(1), S13-S15, (January Supplement).
- Olson, J. R., and Biolski, K. J. (1991). Techniques for Representing Expert Knowledge. In K. A. Ericsson and J. Smith (Eds.), Toward a General Theory of Expertise. New York: Cambridge University Press, Pp. 240-285.
- Schön, D. (1993). The Reflective Practitioner: How Professionals Think in Action. New York: Basic Books, 1993.
- Schvaneveldt, R. W., Durso, F. T., and Dearholt, D. W. (1989). Network Structures in Proximity Data. In G. H. Bower (Ed.), The Psychology of Learning and Motivation. San Diego, CA: Academic Press, Pp. 249-284.
- Schwandt, T. S. (1994). Constructivist, Interpretivist Approaches to Human Inquiry. In N. K. Denzin and Y. S. Lincoln (Eds.), Handbook of Qualitative Research. Thousand Oaks, CA: Sage Publications, Pp. 118-137.

Stevens, R. H. (1991). Search Path Mapping: A Versatile Approach for Visualizing Problem-Solving Behavior. Academic Medicine, 66, S73-S75, (September Supplement).

Stevens, R. H., Lopo, A. C., and Wang, P. (1996). Artificial Neural Networks Can Distinguish Novice and Expert Strategies During Complex Problem Solving. Journal of the American Medical Informatics Association, 3, 131-138.

Wagner, R. K. (1991). Managerial Problem Solving. In R. J. Sternberg and P. A. French (Eds.), Complex Problem Solving: Principles and Mechanisms. Hillsdale, N.J.: Erlbaum, Pp. 159-183.

Table 1

Increase in Similarity of Student and Expert Networks and Similarity Score-Examination Correlations

Expert Referents	Student-Expert Similarity: Mean and Standard Deviation		Similarity Score-Examination Correlations	
	Students		Preinstruction	Postinstruction
	Preinstruction	Postinstruction		
Internists	.22 (.07)	.27 (.07)	.13	.23*
Physiologists	.24 (.08)	.33 (.10)	.13	.27†
Professor	.22 (.08)	.30 (.09)	.17	.27†

* $p < .05$

† $p < .01$

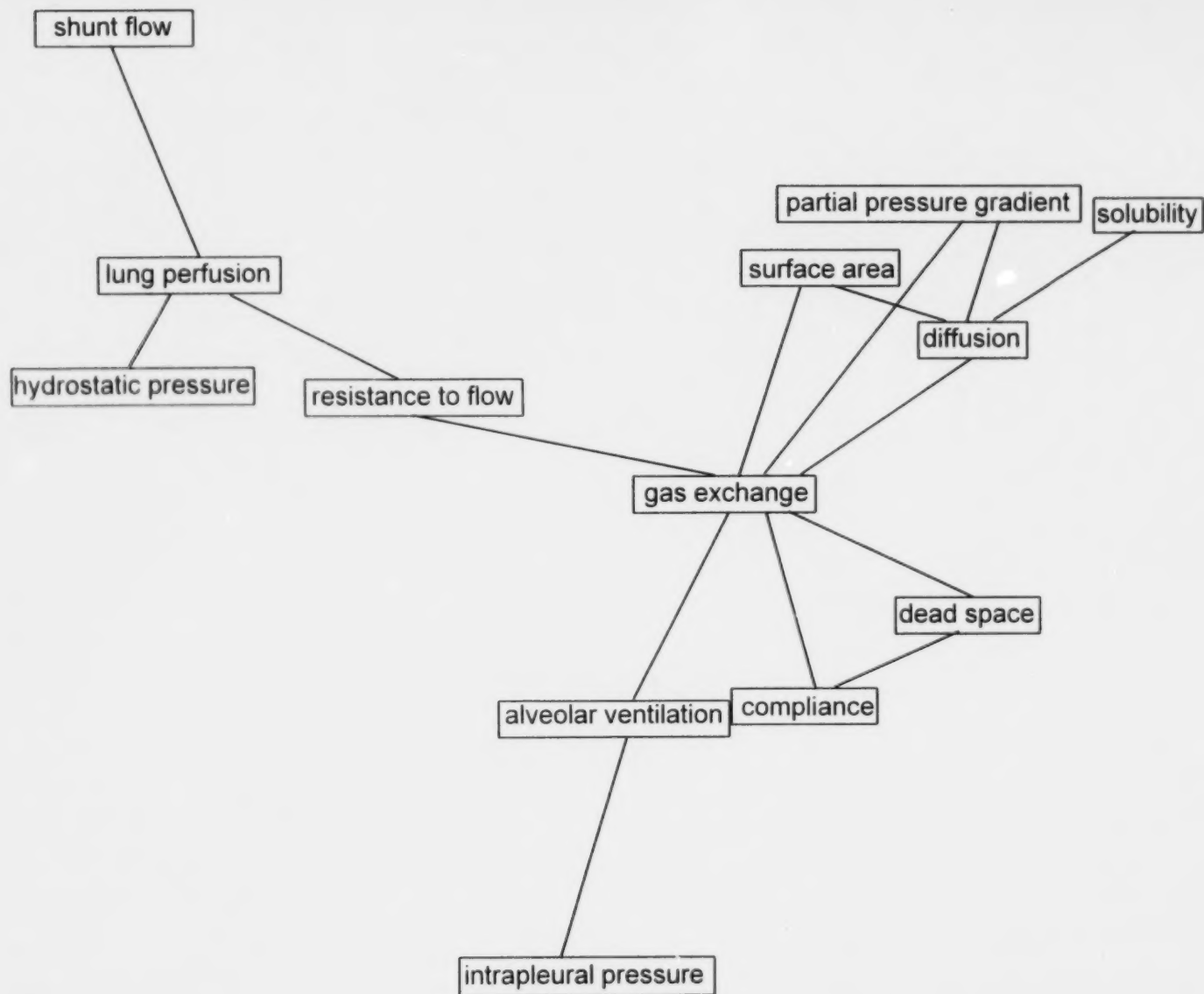


Figure 1
Professor's Pathfinder Network